

## September 28, 2010 – DLC-IT Meeting

### 1. UFDC Server Hardware

- a. UFDC Web Server
  - i. New Shibboleth request for ufdc.ufl.edu
  - ii. www.ufdc.ufl.edu?
- b. Caching Server
  - i. Limited SQL to 4GB (actually 4.3GB?)
  - ii. Cache service still having problems .. debugging
  - iii. Documentation/install for lib-cache-test pending as well
- c. OCR servers
  - i. Prime Recognition
    - 1. OCR2 is working correctly
    - 2. OCR4 on old hardware with a hardware key occasionally hangs requiring Prime restart or full reboot
  - ii. ABBYY
    - 1. Working but needs automation
    - 2. OCR-ing all the image files, not just TIFFs
    - 3. Randall and Maureen working on specifications for automation
    - 4. Will looking at maintenance agreement since a year has almost expired since purchase
- d. DLC Production SAN backup
  - i. Update backup documentation – Logan will send update ( 9/21? )
- e. UFDC server documentation / disaster recovery
  - i. Finishing configuration of complete new instance of UFDC (database copy complete)
  - ii. Caching server pending configuration

### 2. Software Development

#### a. Completed Work

- i. SobekCM/UFDC Web Application ( Version 2.9.6 )
  - 1. Simplified all URLs to move much of the query string data into the URL and utilize URL rewriting to keep the original URL in the user's browser
  - 2. Moved the search stop words into a database table for easy modifications in the future
  - 3. Formalized URL Portals concept
    - a. Now pulled from the database
    - b. Matched by the base url for the request
    - c. Includes default aggregation, default interface, and abbreviation for the library ( i.e., UFDC, dLOC, etc.. )
    - d. Removed all previous hard coding for dLOC/UFDC/etc..
    - e. Added view-only admin mode to see portal details
  - 4. Optimized search queries
    - a. Modified search queries to force SQL to not retain the execution plan for the actual full-text searching subqueries
    - b. Optimized by moving all temporary table creations to the top of query
    - c. Created new basic search query that searches separate table of full citation
    - d. Allowed for new search precisions (exact, contains, stemming, thesaurus look-up)
  - 5. URL rewrite is now used for outputting XML, dataset, or json. No longer forwarded to sobekcm\_data.aspx
  - 6. Online Templates

- a. Corrected main thumbnail element
      - b. Modified templates to include a help button after each element during edit and submission
    - 7. Added ability for system administrators to delete item aggregation aliases online
    - 8. Add ability to browse by metadata lists (i.e., list of all publishers in an item aggregation)
    - 9. Update admin user editing for new user groups ability
    - 10. Added support to export to Excel using new software library
    - 11. Added ability to edit the IP Restrictions through the admin form
    - 12. Add Ability to delete/edit the tags for other users if you have correct rights
    - 13. Corrected case in which AppFabric remote cache manager did not fail gracefully
    - 14. User's language is auto-detected from the HTTP request/browser settings
    - 15. Added internal header for internal users only with ability to search by BibID (more abilities coming later)
  - ii. UFDC Manager ( Version 9.9.9h )
    - 1. New URL format. With the changes, the UFDC Manager was no longer downloading packages
    - 2. Added ability to download the MarcXML. Each marc.xml file is downloaded and a marc\_report.xml is additionally created with all the individual MarcXML files in one file
    - 3. You can now also paste a item URL into the UFDC Manager to download one complete package
  - iii. UFDC Builder ( Version 2.9.6 )
    - 1. Incorporated new code from UFDC web server and navigation changes
  - iv. METS Editor (for internal and state-wide use)
    - 1. Added new elements ( Frequency, Other Titles, Publication Place, Publication Status, FCLA flags )
    - 2. Corrected issue that the DMDID in the struct map was blank
    - 3. Can now view the data as METS and MARC without having to explicitly save first (uses a temp file)
    - 4. Automatically adds TIFFs, JPEGs, JPEG2000s, and text files from the directory to the page images, not just the TIFFs
    - 5. Adds importing from EAD as an option
    - 6. When a user adds a file in a different directory, moves it into the digital resource directory
    - 7. Shows a progress bar / popup when creating checksums as this takes a period of time
    - 8. Converted templates to template schema
    - 9. Added automatic page numbering through the rest of the document
    - 10. Non page-image files appear as second tree view under the structure map tab
  - v. DLC Toolbox ( Version 4.8.5 )
    - 1. Incorporate all of the changes from the METS editor above
    - 2. Updated ImageMagick registry check to include directory check for Win7
    - 3. QC no longer displays packages which only have a thumbs.db file in the folder
- b. Future Pending**
- i. METS Editor
    - 1. Delete a file from the structure map
    - 2. Move multiple pages at once from the structure map
    - 3. Checksum progress bar sometimes doesn't function correctly
    - 4. Save the downloads as a separate structure map in the METS
  - ii. Tracking → UFDC
  - iii. Update dLOC Toolkit
  - iv. Fedora ( no work yet )
  - v. Google map use for entering coordinate information
  - vi. Complete image server / ALTO-aware